

High-Throughput Plasmid cDNA Library Screening

Kenneth H. Wan¹, Charles Yu¹, Reed A. George¹, Joseph W. Carlson¹, Roger A. Hoskins¹, Robert Svirskas², Mark Stapleton¹ and Susan E. Celniker¹

¹ Department of Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA.

² GE Healthcare, 7700 S. River Parkway, Suite #2603, Tempe, AZ 85284

Correspondence should be addressed to S.E. Celniker (celniker@fruitfly.org).
<http://www.fruitfly.org>

ABSTRACT

Libraries of cDNA clones are valuable resources for analysing the expression, structure, and regulation of genes, as well as for studying protein functions and interactions. Full-length cDNA clones provide information about intron and exon structures, splice junctions and 5'- and 3'-untranslated regions (UTRs). Open reading frames (ORFs) derived from cDNA clones can be used to generate constructs allowing expression of native proteins and N- or C-terminally tagged proteins. Thus, obtaining full-length cDNA clones and sequences for most or all genes in an organism is critical for understanding genome functions. Expressed sequence tag (EST) sequencing samples cDNA libraries at random, which is most useful at the beginning of large-scale screening projects. However, as projects progress towards completion, the probability of identifying unique cDNAs via EST sequencing diminishes, resulting in poor recovery of rare transcripts. We describe an adapted, high-throughput protocol intended for recovery of specific, full-length clones from plasmid cDNA libraries in five days.

INTRODUCTION

Traditionally, screens for clones representing a gene of interest were performed by hybridization of labeled, gene-specific, DNA probes to colonies or plaques that were transferred to a nylon filter¹. This and other screening methods, including EST sequencing²⁻⁵; RT-PCR⁶; PCR screening of library pools⁷; RACE⁸; inverse PCR on self-ligated primary double-stranded cDNA⁹⁻¹²; MACH-1 or MACH-2¹³ can be labor and time intensive. This is especially true when the desired clones are rare in the library. Further, these methods are relatively inefficient, yield rearranged or partial transcripts (missing 5'- and/or 3'-UTRs, incomplete ORFs), or fail to identify alternative transcripts.

The basis of this protocol, Self-ligation of Inverse PCR products (SLIP), is a rapid and efficient method for screening plasmid cDNA libraries¹⁴ (Figure 1). The method is inspired by Stratagene's ExSite™ site-directed mutagenesis protocol (<http://www.stratagene.com>), but is intended to recover wild-type clones from a library pool, rather than to mutagenize individual clones. SLIP can yield full-length cDNAs, including 5'- and 3'-UTRs, and can efficiently recover rare transcripts and targeted alternative transcripts. This protocol consists of five steps: PCR; ligation; digestion;

transformation; and DNA sequence verification (Figure 2). The PCR is performed on a circular cDNA library template using 5'-phosphorylated gene-specific primers whose 5' ends abut. The linear PCR products are then self-ligated, using T4 DNA ligase, in a dilute reaction to reconstitute the original cDNA clone. The reaction mix is treated with the restriction enzyme *Dpn* I, which digests the methylated parental DNA plasmids, leaving the un-methylated circular PCR products intact. The plasmids are then transformed into competent cells. Transformants are picked and DNA sequence verified for the correct target.

As is true for any method for isolating full-length cDNAs, SLIP is dependent on starting with high quality, high complexity libraries; it can only recover clones that are present in a cDNA library. To recover clones not in a library, RT-PCR is the standard alternative. Primer design is also critical, as the requirement for abutting primers may limit the number of acceptable primer pairs. Current PCR methodologies impose a practical limit on the size of clones that can be recovered.

In principle, this method can be optimized and applied to any standard plasmid library, including genomic libraries. Although this protocol assumes that all steps are performed in 96-well reaction plates, and that primers have been synthesized in 96-well format, all steps can be scaled up or down to different plate formats or to individual tubes. We have successfully recovered high-quality cDNA clones with inserts of up to 5 kb in our 1.6 kb pOT2 plasmid vector⁵. Adjusting PCR conditions should allow recovery of larger inserts, and advances in PCR techniques should extend the upper size limit of this method. We see little reason why the method should be significantly limited by the lengths of transcripts or cloning vectors.

MATERIALS

Reagents

T4 Polynucleotide Kinase (10U/μl)

New England Biolabs (www.neb.com), Catalog #M0201L

10X T4 Polynucleotide Kinase Buffer

[70mM Tris-Hcl (pH 7.6), 10mM MgCl₂, 5mM dithiothreitol]

Supplied with T4 Polynucleotide Kinase, above.

100X Purified BSA (10mg/mL)

New England Biolabs (www.neb.com), Catalog #B9001S

Adenosine 5' –triphosphate disodium salt

Sigma (www.sigmaaldrich.com), Catalog #A-3377

Sequence-specific primers (20μM)

Invitrogen/Illumina (www.invitrogen.com)

T4 DNA Ligase (400U/μl)

New England Biolabs (www.neb.com), Catalog #M0202L

Dpn I (20U/μl)

New England Biolabs (www.neb.com), Catalog #R0176L

Deoxynucleotide Solution Mix

New England Biolabs (www.neb.com), Catalog #N0447L

Finnzymes Phusion™ High-Fidelity DNA Polymerase (2U/μl)

New England Biolabs (www.neb.com), Catalog #F-530L

5X Phusion™ HF Buffer

Supplied with Phusion™ polymerase, above

TAM-1 Competent Cells

Active Motif (www.activemotif.com), Catalog #11096

LB/Antibiotic Agar Plates

LB Broth Base

Invitrogen (www.invitrogen.com), Catalog #12780-029

Bacto Agar

BD (www.bd.com), Catalog #214010

50mg/ml (stock concentration) antibiotic (Chloramphenicol [Roche (www.roche-applied-science.com)], Catalog #10634433001] is used for the pOT2 vector)

2× YT Media, dehydrated.

VWR Scientific (www.vwrsp.com), Catalog #90003-330

BD Biosciences (www.bdbiosciences.com), Catalog #244020

Plasmid cDNA Library

User supplied

CRITICAL

Successful screening depends on high quality, high complexity libraries; only clones that are present in a cDNA library may be recovered.

Sephadex™ G-50, Superfine

Amersham (www.amershambiosciences.com),

Catalog #17-0041-01

BigDye™ Terminators v3.1 Cycle Sequencing Kit

Applied Biosystems (www.appliedbiosystems.com), P/N 4337455

Sequencing Buffer, 5X

Applied Biosystems (www.appliedbiosystems.com), P/N 4336697

Equipment

100mm x 15mm Petri Dish

BD Falcon (www.bd.com), Catalog #351029

Applied Biosystems 9700 Thermal Cycler

Applied Biosystems (www.appliedbiosystems.com), P/N 4314879

Innova 4300 Incubator Shaker

New Brunswick Scientific (www.nbsc.com), Model #4300

MultiScreen-HV 96-well Plate, 0.45µm, hydrophilic, low protein binding,
Durapore® membrane, clear, non-sterile with lid.

Millipore (www.millipore.com), Catalog #MAHVN4550

MultiScreen Column Loader, 45µl, black, aluminum, non-sterile

Millipore (www.millipore.com), Catalog #MACL09645

MultiScreen Centrifuge Alignment Frame, blue, aqueous applications

Millipore (www.millipore.com), Catalog #MACF09604

MultiScreen Column Loader Scraper

Millipore (www.millipore.com), Catalog #MACL0SC03

MicroAmp Base

Applied Biosystems (www.appliedbiosystems.com), Catalog #N8010531

Soft-face hammer, 1.5 oz

Stanley (www.stanleytools.com), Model #57-592 fitted with Tip #57-582.

NOTE: any small, soft-face hammer will work.

Greiner Bio-One 96-well, U-bottom plate

E&K Scientific (www.eandkscientific.com), Catalog #20101

PCR plate, Prism style optical plate, 96-well, half-skirt

E&K Scientific (www.eandkscientific.com), Catalog #489096

Speedball Roller

E&K Scientific (www.eandkscientific.com), Catalog #EK-5000

PolarSeal Foil Sealing Tape

E&K Scientific (www.eandkscientific.com), Catalog #T592100

SeptraSeal Cap Removal Tool

Matrix Technologies (www.matrixtechcorp.com), Catalog #4469

Ritter Ri-Plate, 2ml, 96-well plates, round bottom

E&K Scientific (www.eandkscientific.com), Catalog #662000

Breathe Seal, breathable sealing tape

E&K Scientific (www.eandkscientific.com), Catalog #T896100

25ml Reagent Reservoir, Divided

Matrix Technologies (www.matrixtechcorp.com), Catalog #8095

REAGENT SETUP

LB/chloramphenicol Agar Plate

20g LB Broth Base, 15g Bacto Agar, 50µg/ml (final concentration) chloramphenicol, 1L purified water.

Approximate yield = 40 plates

- Dissolve LB Broth Base and Bacto Agar in 1L of water.
- Autoclave at 121°C for 30 minutes on liquid cycle.
- Allow solution to cool to approximately 50°C.
- Add 1ml chloramphenicol (50mg/ml), mix well.
- Transfer 25ml of solution into each 100mm x 15 mm Petri dish.
- Cool until solidified, then invert and store at 4°C.

Primer Design

Careful primer design is a critical requirement for success when using this protocol. Although most of the available primer design programs and web interfaces are for standard PCR, an increasing number are incorporating features for alternate PCR techniques, such as degenerate PCR, inverse PCR, and multiplex PCR. A detailed discussion of these programs and primer design theory is beyond the scope of this protocol and has been described in detail¹⁴. Here are some considerations to make when designing primers:

- To recover full-length clones, the primers must abut at their 5' ends, with zero overlap, and in opposite orientation on the template. Beware of even a single base of complementarity at the 5' ends, as this may introduce a point mutation (deletion) in your clone.
- If possible, set the PCR product length parameter of your program equal to the length of the template (including vector).
- Use a mis-priming library to prevent the design of primers complementary to the vector or other species in your library.
- We used the *primer3* (http://www-genome.wi.mit.edu/genome_software/other/primer3.html) software package for primer design. *Primer3* was developed to design primers for standard PCR, so it was necessary to write custom scripts which computationally rearranged the circular template sequences to fit *primer3*'s format requirements. Parameters settings used for *primer3* are as follows¹⁴:

Parameter	Setting
Primer length	23 bases ± 2

Max. number of Ns in primer sequence	0
Product size	Full-length of annotated transcript
T _m	65°C ±5
GC clamp	Most 3' base must be G or C
GC content	50% ±20
Max. complementarity (self)	8*
Max. complementarity (paired primer)	8*
Max. mononucleotide repeat in primer	5 bases
Max. end stability	9*

* See http://www-genome.wi.mit.edu/genome_software/other/primer3.html for score calculation methods.

- Multiple primer pairs were designed for each target and candidate pairs were then subjected to a screening process based on melting temperatures, GC content, and *blastn* (*wublast-2.0*) scores. Values for these parameters were fed into an objective function, and scores were calculated. The primer pairs with the lowest objective function score¹⁴ were chosen for PCR.
- The primer locations within a target gene depend on the parameters listed above, but also on the target itself. For example, to capture full-length cDNAs, primer pairs were restricted to the 5'-most 500 bases of each curated gene model. However, alternative splice variants require the primers be designed within exons not previously captured, which may not be the most 5' exon of the gene model.
- Standard oligonucleotide synthesis produces primers that are not phosphorylated. For the ligation step (Steps 30-35) of this protocol to work, the PCR primers must have a 5' phosphate. Oligonucleotide manufacturers offer 5'-phosphorylation as an additional modification, for an additional cost. Alternatively, the primers may be phosphorylated using T4 polynucleotide kinase, as described in this protocol (Steps 1-7).
- Oligonucleotide synthesis is not perfectly efficient, thus the potential for synthesis errors increases with each base added. Potential errors include base substitutions and truncated oligos. Manufacturers offer purification services at additional cost.

PROCEDURE

Prior to full-scale screening of the cDNA library, perform a smaller-scale screen using established PCR primers. Choose a positive control pair of primers, which targets an ~1.5kb abundant transcript, such as *actin*, or primers that target a transcript known to exist in the library being screened. As the quality and composition of templates within different cDNA libraries vary, adjust PCR parameters (Steps 8-10) for each library, or pool of libraries, until an optimal set of conditions is established for obtaining full-length

targets. The positive control pair can then be used in each subsequent round of the full-scale screen (e.g. designate well H12 as the control reaction for each 96-well plate). Another well can be designated for a negative control.

For the negative control, do not add primers to either the kinase reaction or the PCR; substitute an appropriate volume of water instead. Perform subsequent ligation, digestion, transformation steps as normal.

Kinase Oligo Pairs

- 1| In a reagent reservoir, prepare a reaction mix without primers according to the table below (amounts given are per well). Steps 1-4 are performed on ice.

Kinase Master Mix

Component	Amount (per well)	Amount/concentration in final kinase reaction
10X T4 PNK Buffer	1.5µl	1X
100X Purified BSA	0.15µl	1X
100mM ATP	0.2µl	20 nmol
T4 Polynucleotide Kinase (10U/µl)	1.0µl	10U
dH ₂ O	0.15µl	--
TOTAL	3.0µl	

CRITICAL STEP Phosphorylated primers are required for the subsequent ligation step (Steps 30-35). Alternatively, phosphorylated primers can be purchased from suppliers.

- 2| Aliquot 3µl of reaction mix to each well of a 96-well PCR Plate A.
- 3| Transfer 6µl of 20µM custom forward primers from Primer Plate #1 to Plate A.

CRITICAL STEP Use only high-quality oligonucleotides, as synthesis errors may be incorporated into the amplified product, or generate off-target products.

- 4| Transfer 6µl of 20µM custom reverse primers from Primer Plate #2 to Plate A. Pair the forward and reverse primers (i.e. Well A1 of Primer Plate #1 should pair with Well A1 of Primer Plate #2).
- 5| Seal Plate A with a foil sealer.
- 6| Quick spin Plate A to 1000 rpm in an appropriate centrifuge.
- 7| Place Plate A in a thermal cycler programmed with the following parameters:

Segment	Cycle number	Temperature	Time
1	1	37°C	60 min.
2	2	65°C	20 min.
3	3	4°C	Indefinitely

PAUSE POINT Samples can be left in thermal cycler overnight at 4°C, or stored at -20°C indefinitely.

PCR Setup

- 8| In a clean reagent reservoir, prepare a PCR master mix according to the table below (amounts given are per well). Steps 8-10 are performed on ice.

PCR Master Mix

Component	Amount (per well)	Amount/concentration in final PCR reaction
Plasmid cDNA Library (20-200µg/ml)	1.5µl	2-20ng
5X Phusion Buffer HF	3.0µl	1X
10mM dNTP	0.3µl	3 nmol
Phusion Polymerase Kinase (2U/ µl)	0.15µl	0.3U
dH2O	8.2µl	--
	13.15µl	

CRITICAL STEP The quality of the cDNA library is critical for success. This protocol will recover any species containing the targeted sequences, including genomic contaminants and co-ligated inserts.

Any high-fidelity DNA polymerase may be used ($< 10^{-7}$ error rate). However, to generate blunt ends for the subsequent ligation step (Steps 30-35), choose a polymerase which does not leave an 'A' overhang.

- 9| Aliquot 13.15µl of PCR master mix to a new 96-well PCR Plate B.
- 10| Transfer 1.85µl of kinased primer pairs (Step 7) from Plate A to Plate B.
- 11| Seal Plate B with a foil sealer.
- 12| Quick spin Plate B to 1000 rpm in an appropriate centrifuge.
- 13| Place Plate B in a thermal cycler programmed with the following touchdown PCR¹⁵ cycling parameters:

Segment	Cycle number	Temperature	Time
1	1	98°C	30 seconds
2	2-6	98°C	10 seconds
		72°C --> 68°C, Decrease temp 2°C per cycle	30 seconds
		72°C	2:45 minutes
3	7-36	98°C	10 seconds
		68°C	30 seconds
		72°C	2:45 minutes
4	37	72°C	5 minutes
5	38	4°C	Indefinite

Extension times were determined by manufacturer's suggestion of 30 s per kb for complex templates. Using a 1.6kb vector, we were able to amplify inserts of at least 5kb. Parameters may vary depending on the enzyme you choose. Standard PCR cycling parameters could possibly be used, however we did not try those conditions.

PAUSE POINT Samples can be left in thermal cycler overnight at 4°C, or can be stored at -20°C indefinitely.

Buffer Exchange

- 14| Spread Sephadex G-50 onto MultiScreen Column Loader using the MultiScreen Column Loader Scraper (<http://www.millipore.com/userguides.nsf/docs/p35962>). Perform this step in a secondary container to collect spillover G-50 powder. Spillover can be returned to original bottle and reused.

Filtration through G-50 may reduce concentrations of unincorporated dNTPs and small DNA fragments (< 20 bp). Consult Sephadex manufacturer's literature for size exclusion limits

(<http://www1.amershambiosciences.com/aptrix/upp01077.nsf/Content/Products?OpenDocument&parentid=39901&moduleid=39969>). Alternative buffer exchange methods might be employed. If the T4 DNA ligase and *Dpn* I used are compatible with the buffer used in Step 8, then steps 14-29 can be omitted.

- 15| Scrape the Loader until G-50 powder completely fills the Loader wells and powder level is flush with Loader surface. Visually inspect wells for evenness. Re-apply G-50 if any wells are not flush. Scrape away excess powder from edges of Loader.
- 16| Invert a 96-well, MultiScreen HV plate and place it on the Loader. Align the plate edge with the metal stopper bar.
- 17| Grasp Loader-Plate assembly firmly and invert so the Plate is upright, below the now inverted Loader.
- 18| Tap the Loader gently with the hammer several times along the length of the Loader. Keep the Plate and Loader aligned while doing this.
PAUSE POINT Plates containing dry G-50 powder can be stored indefinitely at room temperature, or until expiration date of G-50.
- 19| Add 300µl dH₂O, per well, to hydrate G-50. Replace MultiScreen HV plate lid.
- 20| Let plate sit at room temp for 3 hours.
PAUSE POINT Hydrated G-50 plates can be stored at 4°C, in an enclosed container with damp paper towels, up to 3 days.
- 21| Place hydrated G-50 plate on top of 96-well, U-bottom microtiter plate fitted with the blue MultiScreen Centrifuge Alignment Frame to collect water.
- 22| Spin at 950 x g for 5 minutes. Discard water from microtiter plate.
- 23| Add additional 150µl dH₂O, per well, to G-50 plate.
- 24| Spin 950 x g for 5 minutes. Discard water from microtiter plate.
- 25| Place G-50 plate on top of a PCR Plate C, no Alignment Frame necessary.
- 26| Place assembly on MicroAmp Base (or any other base that will hold a PCR plate)
- 27| Add 15µl dH₂O to wells of Plate B to bring final volume to 30µl.
- 28| Mix and transfer all 30µl of PCR products onto the G50 columns, be careful to aspirate the sample onto the center of each column. Avoid touching the column.
- 29| Spin at 950 x g for 5 minutes to collect sample in the PCR Plate C.
Approximately ~30µl of sample should be recovered.
PAUSE POINT Samples can be sealed and stored at -20°C.
Up to 15µl of sample may be analysed via agarose gel electrophoresis to confirm presence of PCR products. A fraction could also be saved for further analysis after Step 39.

Ligation

- 30|** In a clean reagent reservoir, prepare a ligation master mix (amounts given are per well). Steps 30-32 are performed on ice.

Ligation Master Mix

Component	Amount (per well)	Amount/concentration in final reaction
10X T4 Ligase Buffer	10µl	1X
T4 DNA Ligase(400U/µl)	1µl	400 U
dH2O	74µl	--
	85µl	

- 31|** Aliquot 85µl of ligation master mix into a new PCR Plate D.
32| Transfer 15µl of gel-filtration purified PCR products from Plate C into Plate D.
33| Seal Plate D with a foil sealer.
34| Quick spin Plate D to 1000 rpm.
35| Place Plate D in a thermal cycler programmed with the following parameters:

Segment	Cycle number	Temperature	Time
1	1-18	16°C	60 minutes
2	19	65°C	10 minutes
3	20	4°C	Indefinite

The 10 minute hold at 65°C inactivates the T4 DNA ligase. We perform an overnight reaction. However, shorter ligation reactions might be used. Consult ligase manufacturer's instructions for details.

PAUSE POINT Samples can be stored at –20°C indefinitely.

Dpn I Digestion

- 36|** Add 1µl *Dpn* I (20U/µl) to each well of Plate D, the ligation reaction. Perform this step on ice.

CRITICAL STEP For the *Dpn* I digestion step, the DNA library must be *dam* methylated before proceeding with this protocol. This should not be an issue if your library underwent a primary amplification in most *E. coli* host strains. If you used a different system to prepare your library, we suggest treating it with *dam* methylase first, although we have not tested the efficacy of this manoeuvre.

- 37|** Seal Plate D with a foil sealer.
38| Quick spin Plate D to 1000 rpm.
39| Place Plate D in a thermal cycler programmed with the following parameters:

Segment	Cycle number	Temperature	Time
1	1-2	37°C	60 minutes

2	3	80°C	20 minutes
3	4	4°C	Indefinite

The 20 minute hold at 80°C inactivates the *Dpn* I restriction enzyme.

PAUSE POINT Samples can be stored at –20°C indefinitely.

Samples may be analysed via agarose gel electrophoresis. If sample still remains from Step 29, both fractions may be run side-by-side for comparison.

Transformation

- 40| Take out tube rack (96 tubes) of TAM-1 chemically competent *E. coli* cells from –80°C freezer
- 41| Remove the plastic base from the competent cell tube rack. This allows consistent heat transfer during subsequent steps.
- 42| Remove the strip caps using the SeptraSeal Cap Removal Tool.
- 43| Thaw TAM-1 cells on wet ice for 5 minutes.
- 44| Add 2µl *Dpn* I digested products, from Plate D, into thawed cells. Use a long-stem pipette tip to reach the bottom of the tube. Tap tubes gently to mix cells.
- 45| Incubate on wet ice for 30 minutes.
- 46| Heat shock cells in 42°C water bath for 30 seconds, gently swirling.
- 47| Place cells back on ice for 2 minutes.
- 48| Add 150µl SOC, per well, for a total volume of 200µl.
- 49| Shake in 37°C incubator shaker, set at 225 rpm for 1 hour.
- 50| Plate 150µl of the transformation reaction onto LB/Chloramphenicol agar plates. Store the remaining outgrowth at 4°C. Grow at 37°C overnight. Next day, for plates with very dense numbers of colonies, plate the remaining 50µl onto LB/Chloramphenicol agar plates.

Our vector has a Cam^R gene, conferring chloramphenicol resistance. Use an appropriate antibiotic for the vector used.

TROUBLESHOOTING

Isolation of DNA from individual clones

- 51| Fill four 2ml deep-well plates with 1.2ml of 2× YT broth, containing chloramphenicol (50 µg/ml) and glycerol (7.5%), per well. Glycerol is used for subsequent freezing of an archival copy of the overnight growth plate
- 52| Pick 4 colonies from plate in step 50 into 2×YT/Chlor/7.5% glycerol blocks (1.2mL each well)
- 53| Seal deep-well blocks with breathable sealers.
- 54| Grow overnight (18-20 hours) at 37°C, 300 rpm.
- 55| Aliquot 100µl of overnight culture into 96-well, U-bottom microtiter plates. This is an archival copy; seal the microtiter plates with foil sealers and store at –80°C.

- 56|** Spin deep-well blocks at 3000 rpm for 5 minutes to pellet the remaining cells. Discard supernatant.
PAUSE POINT Pelleted cells may be stored, sealed, at –20°C for at least three days.
- 57|** Prepare plasmid DNA from the cell pellets. Any standard alkaline-lysis DNA preparation protocol may be used. We have had success using an automated platform (Genomic Solutions RevPrep), as well as manual vacuum-filtration kits (Qiagen QiaPrep 96 Turbo Miniprep, Eppendorf PerfectPrep® 96 Vac Direct Bind).

Sequencing Reactions

- 58|** Set up sequencing reactions, on ice, as follows:

Component	Amount (per well)	Amount/concentration in final reaction
5X Sequencing Buffer	1.75µl	1X*
Water	1.75µl	
Primer (1.6µM)	1.0µl	1.6 pmol
BigDye™ v3.1	0.5µl	1/16 th **
DNA Template	5.0µl	200- 500 ng
TOTAL	10.0µl	

*Buffering is also provided by the BigDye solution.

**The sequencing reactions above are at 1/16th the manufacturer's recommended scale.

Three separate sequencing reactions should be prepared per template:

- 1) Sequence 5' end of insert using a 5' vector-specific primer
- 2) Sequence 3' end of insert using a 3' vector-specific primer
- 3) Sequence internal 5' portion of insert with the custom reverse primer used during PCR (Step 8).

If sequencing cDNAs, the 3' ends should have a poly (A) tail. BigDye™ version 3.1 chemistry does not sequence well through homopolymer stretches.

dRhodamine chemistry sequences better through homopolymer stretches, but is not officially supported on the ABI 3730 series DNA analyser. Alternate chemistries may work.

- 59|** Place sequencing plates on thermal cycler according to manufacturer's recommendations.
- 60|** Remove excess dye-labeled nucleotides according to manufacturer's recommendations.
- 61|** Load sequencing samples onto an automated DNA sequencer and run samples according to manufacturer's recommendations.
- 62|** Analyse sequence data to determine recovery of target clones.

TROUBLESHOOTING

PROCESS TIMELINE

Process bench times are approximate, per 96-well plate (unless otherwise indicated). Some steps may be performed out of sequence, ahead of time, or in parallel to others,

such as Steps 14-19. Time estimates are extrapolated from an experienced technician. Actual times may differ.

Day 1:	Steps 1-6	15 minutes of bench time.
	Step 7	1.5 hour incubation.
	Steps 8-12	15 minutes of bench time.
	Step 13	2.5 hour incubation.
	Steps 14-19	5 minutes of bench time.
	Step 20	3 hour incubation.
	Steps 21-29	20 minutes of bench time.
	Steps 30-34	10 minutes of bench time.
Day 2:	Step 35	18 hour incubation.
	Steps 36-38	5 minutes of bench time.
	Step 39	1.5 hour incubation.
	Steps 40-44	15 minutes of bench time.
	Step 45	30 minute incubation.
	Steps 46-48	5 minutes of bench time.
	Step 49	1 hour incubation.
Day 3:	Step 50	1 hour of bench time, per 96 Petri dishes, using a custom variation of the Copacabana plating method ¹⁶ followed by overnight incubation.
	Steps 51-53	30 minutes of bench time for four (4) 96-well plates.
Day 4:	Step 54	18-20 hour incubation.
	Steps 55-56	30 minutes of bench time for four (4) 96-well plates.
	Step 57	1 hour of bench time, depending on the plasmid preparation method you choose.
	Step 58	15 minutes of bench time.
	Step 59	2.5 hour incubation.
	Step 60	Time depends on protocol used.
	Step 61	30 minutes, depending on instrument used.
Day 5:		Instrument run-times vary.
	Step 62	Time depends on resources available to researcher, number of targets being analysed, etc.

FIGURE LEGENDS AND FIGURES

Figure 1. Overview of the cDNA library screening process by self-ligation of PCR products. Adapted from Nucleic Acids Research, Vol. 33(21) 2005, e185, page 5 Figure 1, with permission of Oxford University Press. (a) Phosphorylated, gene-specific primers for target gene, CG15782, abut at their 5' ends. Primers have zero overlap and are in opposite orientation. (b) Five step process: PCR amplify the target cDNA; self-ligate the linear PCR products; *Dpn* I digest the methylated parental plasmids; transform plasmids into *E. coli* hosts; sequence-verify the transformants.

Figure 2. Five day process flow diagram. Rework loops will add additional days to process time.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

TROUBLESHOOTING TABLE

PROBLEM	SOLUTION
<p>Step 39</p> <p>Agarose gel analysis shows no difference between un-ligated (Step 29) and ligated PCR products (Step 39).</p>	<p>This may indicate that the ligation failed. Repeat ligation. The success of obtaining a clone was not highly correlated with positive agarose gel analysis.</p>
<p>Agarose gel analysis shows multiple bands for PCR products.</p>	<p>Multiple bands may be due to mis-priming OR to alternative splice variants.</p>
<p>Agarose gel analysis shows no bands for ligation reactions.</p>	<p>DNA concentrations in ligation reactions have been greatly diluted. Be sure to account for the dilution when preparing samples for electrophoresis.</p>
<p>STEP 50</p> <p>Few or no colonies.</p>	<p>Verify competency of cells by performing a control transformation according to manufacturer's recommendations.</p> <p>Lack of colonies might also be attributed to problems during the PCR and/or ligation steps. Analyse your ligated (Step 39) and un-ligated PCR products (Step 29) via standard agarose gel electrophoresis¹. Note, in our experience, it is not uncommon to see multiple PCR products per reaction.</p>
<p>Negative control yielded colonies.</p>	<p><i>Dpn</i> I digestion incomplete or failed. Possible cross-contamination of wells, or other contamination.</p>
<p>Experimental wells and negative control have similar numbers of colonies.</p>	<p>See above for negative control. Do not pick colonies for these experimental wells. Repeat digestion and transformation.</p>
<p>Less than 4 colonies in experimental wells and negative control had zero colonies.</p>	<p>Plate remaining transformation outgrowth and/or repeat transformation.</p>
<p>Step 62</p> <p>High proportion of off-target clones.</p>	<p><i>Dpn</i> I digestion (Steps 36-39) incomplete or failed. Repeat Steps 36-39. Make sure your <i>Dpn</i> I enzyme is still active by performing a control reaction on a known</p>

	substrate.
High incidence of primer dimers.	Library complexity may be too low or the clones are not present. Verify your primer design parameters. Adjust primer to template ratio during PCR (Step 8).
High incidence of concatemered primers.	Ligation reaction may not be dilute enough. Try reducing the amount of PCR product in the ligation reaction (Steps 30-32). The goal here is to push the reaction towards intra-molecular ligations rather than inter-molecular ligations. A primer removal step prior to ligation (Steps 30-32) may also help (column purification, alcohol precipitation, phenol extraction, etc.).

ANTICIPATED RESULTS

Results of this protocol are directly correlated with the quality of the cDNA libraries screened. In our pilot study to recover clones of 153 *Drosophila melanogaster* transcription factors, we captured targets represented at levels of less than 1 in 250,000, estimated from a composite of all tissues and stages in our EST sequencing project. Our success rate was 68% in recovery of any gene-specific clone, and 47% in recovery of high quality, full-length clones from four pooled cDNA libraries (adult head; embryo; larva and pupa; and S2 cell line). We found artifacts attributable to the SLIP protocol in 3% of the cases. Artifacts include products with only one of the two primer sequences used for PCR, with multiple, concatenated copies of both primer sequences, and with deletions caused by 5' base complementarity of the primers where they abut. These artifacts are most likely due to primer design or ligation conditions.

ACKNOWLEDGEMENTS

We thank Soo Park, Joshua Karpen, and members of the BDGP for technical support and Gary Karpen for critical reading of the manuscript. This work was supported by National Institutes of Health Grant HG002673 to S.E.C. Author contributions are as follows: R.A.H. conceived the technique and performed the proof-of-concept experiments. M.S., C.Y., and K.H.W. adapted the technique for high-throughput screening, with input from R.A.H, R.A.G and S.E.C. R.A.G and R.S.. wrote custom software scripts to design PCR primers. J.W.C., M.S., and S.E.C. chose targets based on gene predictions. S.E.C. and J.W.C. analysed resulting data from the screens. C.Y. and K.H.W. wrote the high-

throughput protocol, with guidance from R.A.H. and M.S. S.E.C. edited the protocol for publication and provided general supervision.

REFERENCES

1. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* Edn. 2 Vol. 1 1.90 - 1.110, 6.4 - 6.13 (Cold Spring Harbor Laboratory Press, Plainview, NY, 1989).
2. Adams, M. A. et al. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651-1656 (1991).
3. McCombie, W. R. et al. Caenorhabditis elegans expressed sequence tags identify gene families and potential disease gene homologues. *Nat Genet* **1**, 124-31 (1992).
4. Delseny, M., Cooke, R., Raynal, M. & Grellet, F. The Arabidopsis thaliana cDNA sequencing projects. *FEBS Lett* **405**, 129-32 (1997).
5. Rubin, G. M. et al. A Drosophila complementary DNA resource. *Science* **287**, 2222-4 (2000).
6. Mocharla, H., Mocharla, R. & Hodes, M. E. Coupled reverse transcription-polymerase chain reaction (RT-PCR) as a sensitive and rapid method for isozyme genotyping. *Gene* **93**, 271-5 (1990).
7. Munroe, D. J. et al. Systematic screening of an arrayed cDNA library by PCR. *Proc Natl Acad Sci U S A* **92**, 2209-13 (1995).
8. Frohman, M. A., Dush, M. K. & Martin, G. R. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci U S A* **85**, 8998-9002 (1988).
9. Huang, S. H., Chen, S. H. & Jong, A. Y. Use of inverse PCR to clone cDNA ends. *Methods Mol Biol* **221**, 51-8 (2003).
10. Ochman, H., Gerber, A. S. & Hartl, D. L. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621-3 (1988).
11. Triglia, T., Peterson, M. G. & Kemp, D. J. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res* **16**, 8186 (1988).
12. Green, I. R. & Sargan, D. R. Sequence of the cDNA encoding ovine tumor necrosis factor-alpha: problems with cloning by inverse PCR. *Gene* **109**, 203-10 (1991).
13. Haerry, T. E. & O'Connor, M. B. Isolation of Drosophila activin and follistatin cDNAs using novel MACH amplification protocols. *Gene* **291**, 85-93 (2002).
14. Hoskins, R. A. et al. Rapid and efficient cDNA library screening by self-ligation of inverse PCR products (SLIP). *Nucleic Acids Res* **33**, e185 (2005).
15. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K. & Mattick, J. S. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* **19**, 4008 (1991).
16. Worthington, M. T., Luo, R. Q. & Pelo, J. Copacabana method for spreading E. coli and yeast colonies. *Biotechniques* **30**, 738-40, 742 (2001).

a**Target CG15782**Gene-specific forward primer
(phosphorylated)

P 5' TGTCTATGCCCCGCGAAATGCTC 3'
 5' TGGAGGAGGCCTTCAAGGAAGCCCATTATCCCGATGTCTATGCCCCGCGAAATGCTCTCGCTGAAGA 3'
 3' ACCTCCTCCGGAAGTTCTTCGGGTAATAGGGCTACAGATACGGGCGCTTTACGAGAGCGACTTCT 5'
 3' GTTCCTTCGGGTAATAGGGCT 5' P

Gene-specific reverse primer
(phosphorylated)**b**

